UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/565,417	08/03/2006	William J. Welsh	UMD0067US.NP	2667

46046 7590 11/21/2008

LICATA & TYRRELL P.C.
66 EAST MAIN STREET
MARLTON, NJ 08053

EXAMINER	
SMITH, CAROLYN L	
ART UNIT	**PAPER NUMBER**
1631	

NOTIFICATION DATE	DELIVERY MODE
11/21/2008	ELECTRONIC

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Notice of the Office communication was sent electronically on above-indicated "Notification Date" to the following e-mail address(es):

poreilly@licataandtyrrell.com

Office Action Summary	Application No. 10/565,417	Applicant(s) WELSH ET AL.	
	Examiner Carolyn Smith	Art Unit 1631	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE _3_ MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on _03 September 2008_.

2a)☒ This action is **FINAL**. 2b)☐ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) _1 and 3_ is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) _1,3_ is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☐ The drawing(s) filed on _____ is/are: a)☐ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☐ Information Disclosure Statement(s) (PTO/SB/08) Paper No(s)/Mail Date _____.
4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____ .
5) ☐ Notice of Informal Patent Application
6) ☐ Other: _____.

DETAILED ACTION

Applicant's amendments and remarks, filed 9/3/08, are acknowledged. Amended claim 1 and cancelled claims 2 and 4-8 are acknowledged.

Applicant's arguments, filed 9/3/08, have been fully considered but they are not deemed to be persuasive. Rejections and/or objections not reiterated from the previous office actions are hereby withdrawn. The following rejections and/or objections are either reiterated or newly applied. They constitute the complete set presently being applied to the instant application.

Claims 1 and 3 are herein under examination.

Claim Rejections - 35 USC § 103

The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

Claims 1 and 3 are rejected under 35 U.S.C. 103(a) as being unpatentable over Troyanskaya et al. (Bioinformatics, 2001, Volume 17, Number 6, pages 520-525) in view of Cunningham (US 2002/0129038 A1) and Xu et al. (US 2006/0241923 A1) with additional support from online Merriam-Webster dictionary ("Gaussian" definition). This rejection is necessitated by amendment.

Troyanskaya et al. describe methods for estimating missing values in DNA microarrays via imputing (abstract and title). Troyanskaya et al. describe k-means clustering and various model-based approaches and algorithms, such as (Single Value Decomposition) SVDimpute algorithm via normalization for microarray data comprising rows and columns (page 520, col. 2, first and second paragraphs; page 521, col. 1, first and second and fourth paragraphs and col. 2, first and last paragraph). According to the online Merriam-Webster dictionary, the definition of "Gaussian" is "being or having the shape of a normal curve or a normal distribution" (this definition is not being used as prior art, but rather to clarify the definition of the term "Gaussian"). The normalization of data represents normal distributions or Gaussian distributions or models. Troyanskaya et al. describe using k eigengenes, using a row average, and an expectation maximization method that is repeated until the change falls below a threshold (converges) (page 522, col. 1, third and fourth paragraphs). Troyanskaya et al. describe a website, software and methods implemented on a computer (abstract and page 524, col. 1, last paragraph) which represents a computer readable medium and program and computer which inherently contains memory and output of missing values. Troyanskaya et al. do not recite a model which imposes a mixture of multivariate normal distributions or using Bayesian information criterion.

Cunningham describes a computer system and computer readable media with data storage devices for using an algorithm and improvements to it (0081) including dealing with missing values and inserting estimated values for C, R, and W matrices and estimates log-likelihood to obtain global means (0047-0048, 0085-0086, Figure 2A (206), 0056, 0136), storing values for each data point (0052), clustering data by Gaussian mixture clustering by imposing a

mixture of multivariate normal distributions (abstract, 0015, 0042), including determining the value of K (number of clusters) (Table 1), partitioning of data (0016, 0032, 0114), repeating a expectation-maximization algorithm until convergence (claim 2; 0047, 0093-0094, 0128-0129) that is performed in a computer implemented data mining system to create a Gaussian Mixture Model as well as generating output storing probabilities for each point belonging to each of the clusters (0056) and describing clustering in the data by computing a mixture of multivariate normal distributions (abstract, 0015, 0028, 0030-0033, 0042). Cunningham does not describe using Bayesian information criterion.

Xu et al. describe imputing missing values in data (0036) and building statistical models by clustering data (0037, 0048) by using Bayesian information criteria (0040), and imputing using a mean value (0061)(i.e. estimating missing values).

It would have been obvious to one of ordinary skill in the art at the time the invention was made to use the expectation-maximization method computing a mixture of multivariate normal distributions of Cunningham in the method of Troyanskaya et al. wherein the motivation would have been to employ a clustering algorithm that can work with large datasets and provide significant enhancements to a Gaussian Mixture Model, as stated by Cunningham (0015-0016, 0022) since there is a need to increase the range of data sets to which the algorithms can be applied, as stated by Troyanskaya et al. (abstract). It would have been further obvious to impute missing values via a model involving Bayesian information criteria as taught by Xu et al. in the methods of Troyanskaya et al. and Cunningham wherein the motivation would have been to generate statistical models more quickly and with better quality via an automated approach to adopt new strategies more rapidly, as stated by Xu et al. (0008).

Thus, Troyanskaya et al. in view of Cunningham and Xu et al. make obvious the invention.

Claim Rejections - 35 USC § 103

The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

Claims 1 and 3 are rejected under 35 U.S.C. 103(a) as being unpatentable over Hytopoulos et al. (US 2002/0169560 A1) in view of Cunningham (US 2002/0129038 A1) and Xu et al. (US 2006/0241923 A1) with additional support from online Merriam-Webster dictionary ("Gaussian" definition). This rejection is necessitated by amendment.

Hytopoulos et al. describe a computer-implemented method and a system using microarray expression data arrays, cluster arrays, and clustering tools wherein the expression values have been normalized, filtered, and imputed, wherein missing data are imputed, and outputted (abstract and paragraphs 0002, 0052, 0084, and 0123). According to the online Merriam-Webster dictionary, the definition of "Gaussian" is "being or having the shape of a normal curve or a normal distribution" (this definition is not being used as prior art, but rather to clarify the definition of the term "Gaussian"). The normalization of data represents normal distributions or Gaussian distributions or models. Hytopoulos et al. describe using a computer readable medium in association with a computer including a processor and memory and

computer instructions which are configured to cause a computer to process data (claim 15) which represents an algorithm and computer software program and product. Hytopoulos et al. describe allowing the user to select K-nearest neighbor imputation mechanism or other data imputation mechanisms (paragraph 0125). Hytopoulos et al. describe analysis of gene expression data to form clusters (abstract). Hytopoulos et al. describe identifying genes represented in respective rows (paragraph 0038) which represents a partitioning of rows of microarray data. Hytopoulos et al. describe mapping rows of expression data (paragraph 0131). Hytopoulos et al. do not describe a model which imposes a mixture of multivariate normal distributions or using Bayesian information criterion.

Cunningham describes a computer system and computer readable media with data storage devices for using an algorithm and improvements to it (0081) including inputting data from a set of points, dealing with missing values and inserting estimated values for C, R, and W matrices and estimates log-likelihood to obtain global means (0047-0048, 0085-0086, Figure 2A (200 and 206), 0055-0056, 0136), storing values for each data point (0052), clustering data by Gaussian mixture clustering by imposing a mixture of multivariate normal distributions (abstract, 0015, 0042), including determining the value of K (number of clusters) (Table 1), partitioning of data (0016, 0032, 0114), repeating a expectation-maximization algorithm until convergence (claim 2; 0047, 0093-0094, 0128-0129) that is performed in a computer implemented data mining system to create a Gaussian Mixture Model as well as generating output storing probabilities for each point belonging to each of the clusters (0056) and describing clustering in the data by computing a mixture of multivariate normal distributions (abstract,

0015, 0028, 0030-0033, 0042). Cunningham does not describe using Bayesian information criterion.

Xu et al. describe imputing missing values in data (0036) and building statistical models by clustering data (0037, 0048) by using Bayesian information criteria (0040), and imputing using a mean value (0061)(i.e. estimating missing values).

It would have been obvious to one of ordinary skill in the art at the time the invention was made to use the expectation-maximization method computing a mixture of multivariate normal distributions of Cunningham in the method of Hytopoulos et al. wherein the motivation would have been to employ a clustering algorithm that can work with large datasets and provide significant enhancements to a Gaussian Mixture Model, as stated by Cunningham (0015-0016, 0022) since the amount of genetic data is quite large and an effective mechanism is needed to determine which genes are correlated with various human conditions, as stated by Hytopoulos et al. (0004 and 0009). It would have been further obvious to impute missing values via a model involving Bayesian information criteria as taught by Xu et al. in the methods of Hytopoulos et al. and Cunningham wherein the motivation would have been to generate statistical models more quickly and with better quality via an automated approach to adopt new strategies more rapidly, as stated by Xu et al. (0008).

Thus, Hytopoulos et al. in view of Cunningham and Xu et al. make obvious claims 1 and 3.

Claim Rejections – 35 USC § 103

The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

This application currently names joint inventors. In considering patentability of the claims under 35 U.S.C. 103(a), the examiner presumes the subject matter of the various claims was commonly owned at the time any inventions covered therein were made absent any evidence to the contrary. Applicant is advised of the obligation under 37 CFR 1.56 to point the inventor and invention dates of each claim that was not commonly owned at the time a later invention was made in order for the examiner to consider the applicability of 35 U.S.C. 103(c) and potential 35 U.S.C. (e), (f) or (g) prior art under 35 U.S.C. 103(a).

Claims 1 and 3 are rejected under 35 U.S.C. 103(a) as being unpatentable over Hytopoulos et al. (US2002/0169560 A1) with additional support from online Merriam-Webster dictionary ("Gaussian" definition) in view of Cereghini et al. (US 6,496,834 B1) and Xu et al. (US 2006/0241923 A1). This rejection is necessitated by amendment.

Hytopoulos et al. describe a computer-implemented method and a system using microarray expression data arrays, cluster arrays, and clustering tools wherein the expression values have been normalized, filtered, and imputed, wherein missing data are imputed, and outputted (abstract and paragraphs 0002, 0052, 0084, and 0123). According to the online Merriam-Webster dictionary, the definition of "Gaussian" is "being or having the shape of a

normal curve or a normal distribution" (this definition is not being used as prior art, but rather to clarify the definition of the term "Gaussian"). The normalization of data represents normal distributions or Gaussian distributions or models. Hytopoulos et al. describe using a computer readable medium in association with a computer including a processor and memory and computer instructions which are configured to cause a computer to process data (claim 15) which represents an algorithm and computer software program and product. Hytopoulos et al. describe allowing the user to select K-nearest neighbor imputation mechanism or other data imputation mechanisms (paragraph 0125). Hytopoulos et al. describe analysis of gene expression data to form clusters (abstract). Hytopoulos et al. describe identifying genes represented in respective rows (paragraph 0038) which represents a partitioning of rows of microarray data. Hytopoulos et al. describe mapping rows of expression data (paragraph 0131). Hytopoulos et al. do not describe repeating a classification expectation-maximization algorithm until the K partitions converge or a model which imposes a mixture of multivariate normal distributions or using Bayesian information criterion.

Cereghini et al. describe a method of performing cluster analysis inside a relational database management system using Gaussian mixture parameters and implementing an Expectation-Maximization (EM) clustering algorithm iteratively (abstract). Cereghini et al. describe grouping a set of data into k clusters with k rows (partitioned) (col. 2, lines 57-63). Cereghini et al. describe the expectation-maximization algorithm converges quickly and performing iterations (col. 9, lines 34-42). Cereghini et al. describe the EM algorithm assumes the data is formed by the mixture of multivariate normal distributions. Cerehini et al. do not describe using Bayesian information criterion.

Xu et al. describe imputing missing values in data (0036) and building statistical models by clustering data (0037, 0048) by using Bayesian information criteria (0040), and imputing using a mean value (0061)(i.e. estimating missing values).

Hytopoulos et al. state that effective mechanisms for analyzing DNA array data are needed to determine which genes or combination of genes are correlated to various human conditions (paragraph 0009). Cereghini et al. state the EM algorithm is robust for noisy data and missing information (col. 7, lines 5-6). Cereghini et al. state cluster analysis does not typically work well with large databases due to memory limitations and the execution times required (col. 2, lines 32-39). It would have been obvious to one of ordinary skill in the art at the time the invention was made to use effective means for analyzing DNA array data, as stated by Hytopoulos et al., by using algorithms supporting large databases, as stated by Cereghini et al. The person of ordinary skill in the art would have been motivated to make that modification in order to find effective ways (as stated by Hytopoulos et al. and Cereghini et al.) of correlating genes to human conditions (as stated by Hytopoulos et al.) by allowing non-statisticians to benefit from advanced mathematical techniques available in a relational environment, as stated by Cereghini et al. (col. 2, lines 40-43). It would have been further obvious to impute missing values via a model involving Bayesian information criteria as taught by Xu et al. in the methods of Hytopoulos et al. and Cereghini et al. wherein the motivation would have been to generate statistical models more quickly and with better quality via an automated approach to adopt new strategies more rapidly, as stated by Xu et al. (0008).

Thus, Hytopoulos et al. with additional support from the online Merriam-Webster dictionary, in view of Cereghini et al. and Xu et al. make obvious the instant invention.

Applicant summarizes each rejection. Applicant argues that the rejections do not describe where the prior art teaches estimating missing values by a GMCimpute algorithm and refer to the paragraph spanning pages 17 and 18 of the specification which states the algorithm takes the average of all K_estimates by S components, wherein each missing entry has S estimates and the final estimate is the average of them. It is the noted that while claims can be read in light of the specification, limitations in the specification cannot be read into the claims. It is noted that not all limitations need to come from a single reference in a 35 USC 103 rejection. Troyanskaya et al. describe missing value estimation methods of microarrays using various algorithms (title and abstract) and Cunningham describes using an EM algorithm and improvements to it (0081) including dealing with missing values and inserting estimated values for C, R, and W matrices and estimates log-likelihood to obtain global means (0047-0048, 0085-0086, Figure 2A (206), 0056, 0136). Hytopoulos et al. describe a computer-implemented method and a system using microarray expression data arrays, cluster arrays, and clustering tools wherein the expression values have been normalized, filtered, and imputed, wherein missing data are imputed, and outputted (abstract and paragraphs 0002, 0052, 0084, and 0123). Xu et al. describe imputing missing values in data (0036) and building statistical models by clustering data (0037, 0048) by using Bayesian information criteria (0040), and imputing using a mean value (0061)(i.e. estimating missing values). Applicant argues that the prior art fails to teach the use of Bayesian information criterion to select the number of clusters in the Gaussian mixture clustering as stated in the specification on pages 13 and 14. It is again noted that while claims can be read in light of the specification, limitations in the specification cannot be read into the

claims. Xu et al. describe imputing missing values in data (0036) and building statistical models by clustering data (0037, 0048) by using Bayesian information criteria (0040), and imputing using a mean value (0061)(i.e. estimating missing values). In addition, for example, Cunningham describes a computer system and computer readable media with data storage devices for using an algorithm and improvements to it (0081) including dealing with missing values and inserting estimated values for C, R, and W matrices and estimates log-likelihood to obtain global means (0047-0048, 0085-0086, Figure 2A (206), 0056, 0136), storing values for each data point (0052), clustering data by Gaussian mixture clustering by imposing a mixture of multivariate normal distributions (abstract, 0015, 0042), including determining the value of K (number of clusters) (Table 1), partitioning of data (0016, 0032, 0114). Applicant is reminded that not all limitations need to be taught in a single reference for a 35 USC 103 rejection. Applicant summarizes the framework for objective analysis for determining obviousness under 35 USC 103. Applicant argues that the prior art fails to teach every limitation in the instant claims. This statement is found unpersuasive as each limitation is addressed, as fully described in each rejection above.

Other prior art made of record

Although not being used as prior art, Yeung et al.'s "Model-based clustering and data transformations for gene expression data" (Bioinformatics, 2001, Volume 17, Number 10, pages 977-987) is being put on the record. Yeung et al. discuss Gaussian mixture models for clustering in gene expression data analysis and the ability to incorporate missing data into the model.

Conclusion

No claim is allowed.

Applicant's amendment necessitated the new ground(s) of rejection presented in this Office action. Accordingly, **THIS ACTION IS MADE FINAL**. See MPEP § 706.07(a). Applicant is reminded of the extension of time policy as set forth in 37 CFR 1.136(a).

A shortened statutory period for reply to this final action is set to expire THREE MONTHS from the mailing date of this action. In the event a first reply is filed within TWO MONTHS of the mailing date of this final action and the advisory action is not mailed until after the end of the THREE-MONTH shortened statutory period, then the shortened statutory period will expire on the date the advisory action is mailed, and any extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of the advisory action. In no event, however, will the statutory period for reply expire later than SIX MONTHS from the date of this final action.

Papers related to this application may be submitted to Technical Center 1600 by facsimile transmission. Papers should be faxed to Technical Center 1600 via the PTO Fax Center. The faxing of such papers must conform with the notices published in the Official Gazette, 1096 OG 30 (November 15, 1988), 1156 OG 61 (November 16, 1993), and 1157 OG 94 (December 28, 1993) (See 37 CFR §1.6(d)). The Central Fax Center number for official correspondence is (571) 273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished

applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. If you have questions on access to the Private PAIR system, please contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, please call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Carolyn Smith, whose telephone number is (571) 272-0721. The examiner can normally be reached Monday through Thursday from 8 A.M. to 6:30 P.M.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Marjorie Moran, can be reached on (571) 272-0720.

November 13, 2008

/Carolyn Smith/
Primary Examiner
AU 1631